

---

# **Design and implementation of semantic search engine Smartch**

---

**Kunmei Wen, Zhengding Lu, [Ruixuan Li](#), Xiaolin Sun**

**Intelligence and Distributed Computing Lab**

**College of Computer Science and Technology**

**Huazhong University of Science and Technology**

---

# Semantic search

- Categorize

- Incremental semantic search engine based on traditional search engine[3-6]
- Intelligent semantic search based on ontology inference [7-10]
- Other semantic search [11-13]

- Questions

- Complicated queries?
- User interface?

---

# Outline

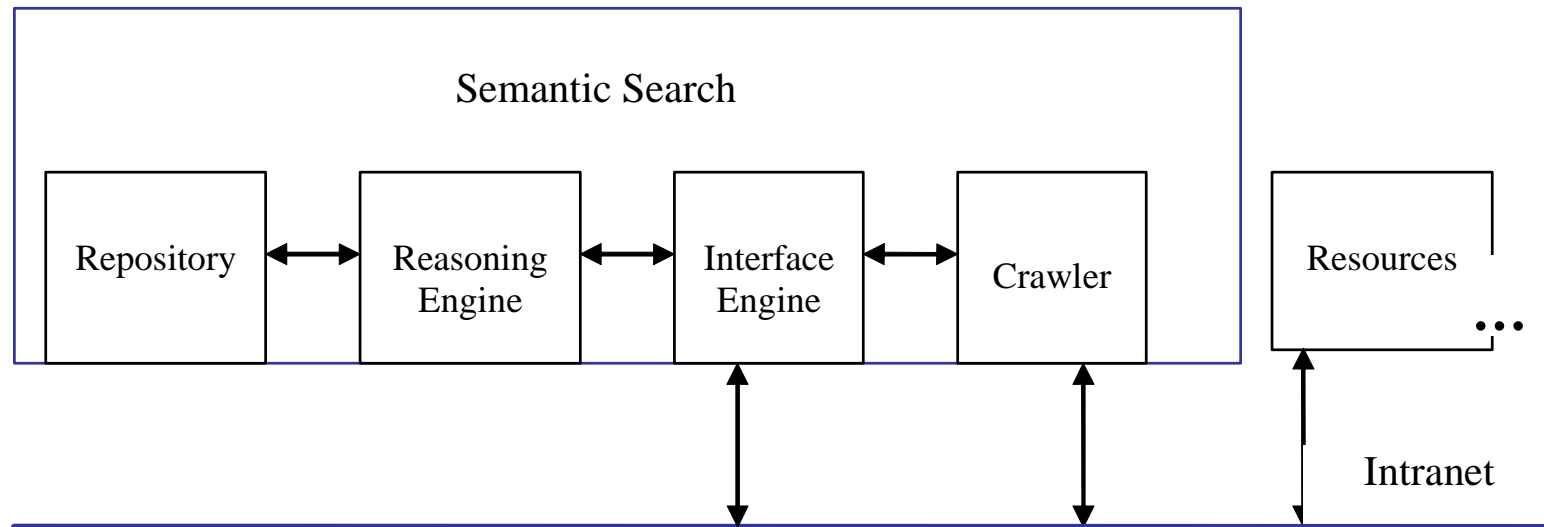
- **Architecture of Semantic Search**
- **System Implementation**
- **Experiment and Conclusions**

---

# Outline

- **Architecture of Semantic Search**
- **System Implementation**
- **Experiment and Conclusions**

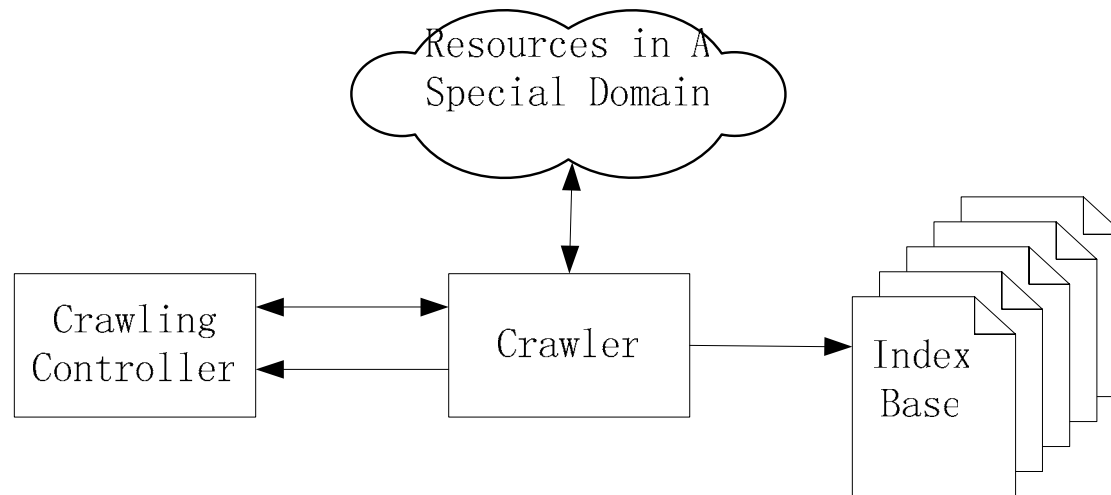
# Architecture of Semantic Search



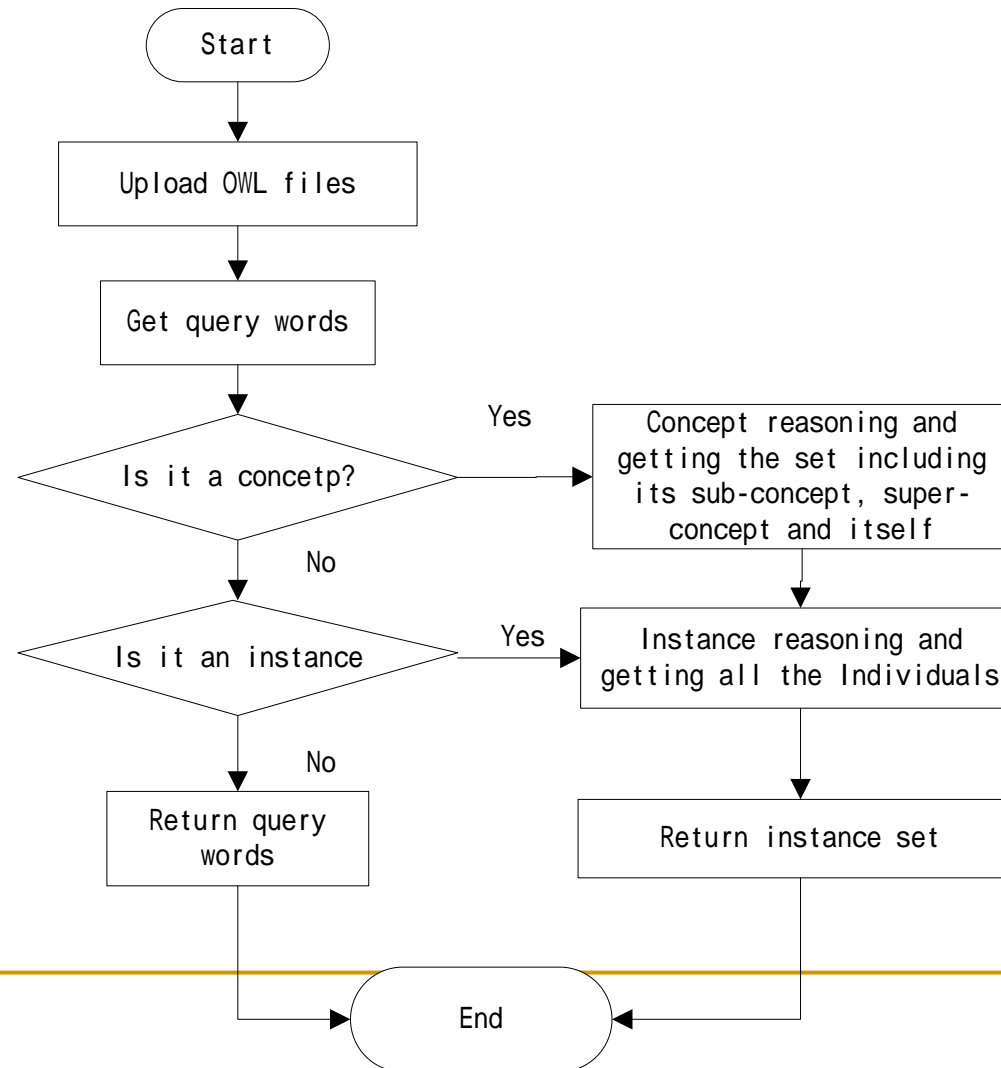
- 
- Repository based on ontology: Repository stores domain ontology owl files
  - Crawler: Crawler's task is catching the intranet resources and building index for them
  - Reasoning engine: Reasoning service of semantic search is carried out by reasoning engine
  - Interface engine: User can interface with semantic search engine

---

# Crawling process of semantic search engine

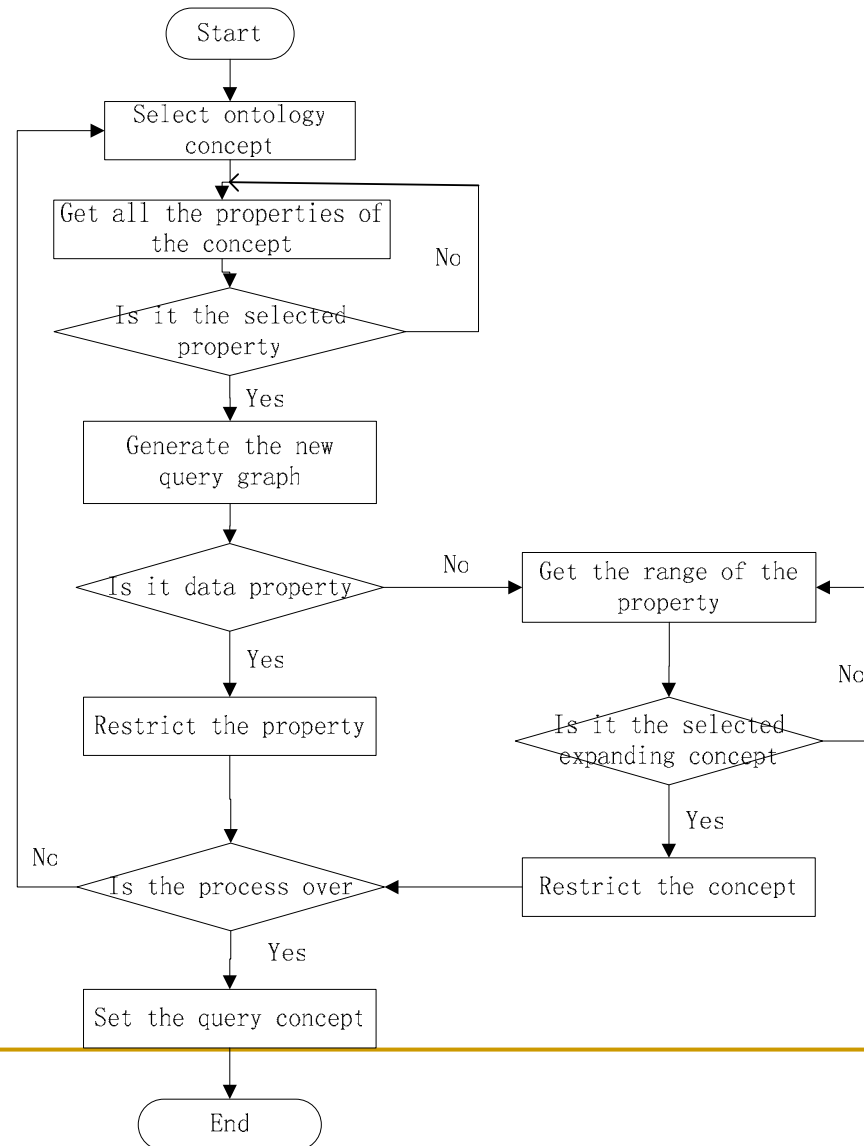


# Reasoning process of semantic search engine





# User-defined process of graphic query



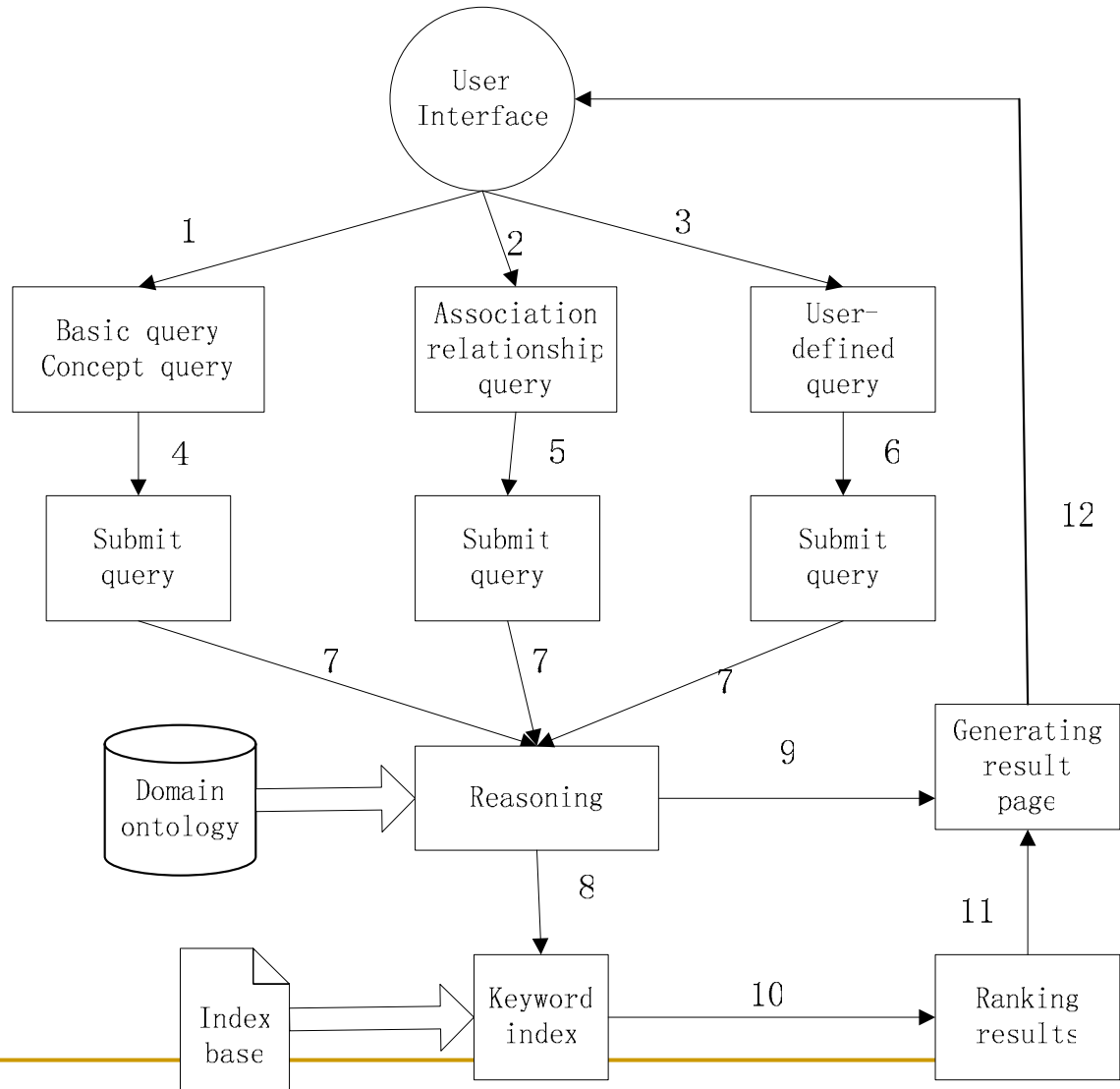
---

# Outline

- **Architecture of Semantic Search**
- **System Implementation**
- **Experiment and Conclusions**

- 
- Lucence: Intranet search tool
  - Jena: ontology parsing tool
  - SQLServer: ontology data persistence
  - SVG、Grahpviz: User-defined process of graphic query

# Whole process of semantic search



---

# Outline

- **Architecture of Semantic Search**
- **System Implementation**
- **Experiment and Conclusions**

# Basic search

SmartCH

搜索

[返回首页](#) [登录](#)

共有26项符合 **论文** 的查询结果 (搜索用时: 0.032秒)

[论文发表](#)

| [论文发表](#) | [论文发表最新论文](#) | [已发表论文](#) ...

<http://idc.hust.edu.cn/~rxli/chinese/publications.htm>

[Oracle Developer 2000 学习教程](#)

[论文发表](#)...

<http://idc.hust.edu.cn/~rxli/chinese/oracle2000.htm>

[Oracle Developer 2000 学习教程](#)

[论文发表](#)..

<http://idc.hust.edu.cn/~rxli/chinese/books.htm>

[Some useful notes on how to prepare and deliver a good presentation](#)

... You will present a **paper** inequally to all **papers** and to all speaker... de some guidance in **paper**to each **paper** to help the present...lved in preparing a **paper** for presentation. I... and understand t...

# Concept search

SmartCH

搜索

[返回首页](#)

共有23项符合 **教师** 的查询结果 (搜索用时: 0.093秒)

## 《李瑞轩》

[Oracle Developer 2000 学习教程](#)

译作者: **李瑞轩**, 杨曼红, 王俊 ...

<http://idc.hust.edu.cn/~rxli/chinese/books.htm>

## 《文坤梅》

[论文发表](#)

**李瑞轩**, **卢正鼎**. **李瑞轩李瑞轩**, **卢正鼎**, 於光灿. P2P环境下基于多维数据的资源78. 於光灿, **卢正鼎**, **李瑞轩李瑞轩**, 赵战西, 王治纲, **文坤梅**, 孙小林. 基于本体的多域访问控制策略集**李瑞轩**75. 於光灿, **卢正鼎**, **李瑞轩**74. 於光灿, **卢正鼎**, **李瑞轩**73. 宋伟, **卢正鼎**, **李瑞轩**59. 彭俊杰, **卢正鼎**, **李瑞轩**56. 唐卓, **卢正鼎**, **李瑞轩**55. 胡和平, 黄保华, 姚寒冰, **卢正鼎**, **李瑞轩李瑞轩李瑞轩**, 赵战西, 王治...

<http://idc.hust.edu.cn/~rxli/chinese/publications.htm>

发表著作: OracleDeveloper  
主讲: WebPeer  
发表著作: 一种改进的非结构  
指导的学生: TerefeMuleta  
发表著作: paper004  
授课: 数据库系统分析与设计  
指导的学生: 吉勇  
授课: 语义网与本体论  
主讲: 对等计算技术的应用及  
姓名: 李瑞轩  
授课: 数据库系统原理  
研究方向: 系统安全; 分布式  
发表著作: paper003  
主讲: 斯坦福大学数据库研究  
发表著作: 基于P2P的协同设  
发表著作: paper005  
研究方向: 分布计算; 分布式  
指导的学生: 赵战西  
指导的学生: 高昶  
指导的学生: 章华娟  
授课: 对等计算  
发表著作: 一种Peer-to-Peer  
发表著作: 多数据库系统原理  
指导的学生: 胡劲伟  
发表著作: 在Java2环境中实  
籍贯: 湖北  
授课: 面向对象程序设计

# User-defined graphics mode

The screenshot displays the SmartCH web application interface. At the top, there are two browser tabs labeled "Smartch". In the top right corner, there are navigation links: "基本搜索" (Basic Search), "概念搜索" (Concept Search), "相关帮助" (Related Help), and "返回首页" (Return Home). The main heading "SmartCH" is prominently displayed in the center, with a subtitle below it: "-----基于SVG图形化定制查询" (-----Based on SVG Graphical Customization Query).

Below the heading, there is a control bar with three buttons: "新的查询" (New Query) with a "new" input field, "清空查询" (Clear Query) with a "clear" input field, and "递交查询" (Submit Query) with a "query" input field.

The main content area is divided into two sections. On the left, a "提示 tips:" (Tips) box contains a list of instructions:

1. 从类别导航树中选取某个类别;
2. 针对所有类别的“关联到其它类别”可以建立类别间的关系;
3. 针对查询类别的“添加属性限制”从属性上约束查询类;
4. 对于非查询类别“类别实例化”可以输入实例来限定该类别;
5. “仅保留该类别”将删除选中类别外所有类别和联系;

On the right, there is a tree view of categories, each with a radio button and a plus sign:

- ☐ 人员
- ☐ 研究领域
- ☐ 会议
- ☐ 著作
- ☐ 实验室
- ☐ 设备
- ☐ 报告
- ☐ 课程
- ☐ 项目

Below the tree view, there is a text prompt: "单击某类别后 进入查询定制过程..." (Click a category to enter the query customization process...).



# User-defined graphics mode

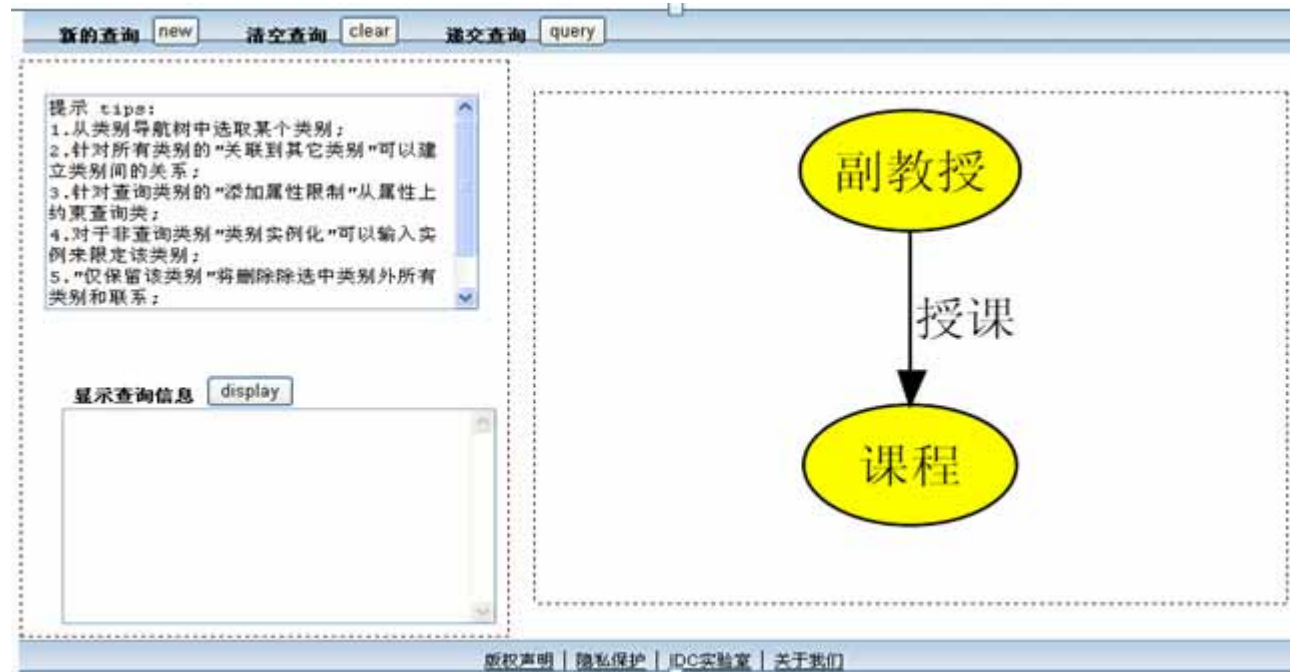
新的查询  清空查询  递交查询

提示 tips:  
1. 从类别导航树中选取某个类别;  
2. 针对所有类别的“关联到其它类别”可以建立类别间的关系;  
3. 针对查询类别的“添加属性限制”从属性上约束查询类;  
4. 对于非查询类别“类别实例化”可以输入实例来限定该类别;  
5. “仅保留该类别”将删除选中类别外所有类别和联系;

显示查询信息

- 人员
  - 其它
  - 学生
    - 博士生
    - 硕士
    - 硕士生
    - 博士
  - 教师
    - 教授
    - 讲师
    - 副教授
    - 助教
- 研究领域
  - 一级研究领域
  - 二级研究领域
- 会议
- 著作
  - 书籍
  - 论文
- 实验室
- 设备
  - 投影仪

# User-defined graphics mode



# User-defined graphics mode

新的查询  清空查询  递交查询

提示 tips:  
1. 从类别导航树中选取某个类别;  
2. 针对所有类别的“关联到其它类别”可以建立类别间的关系;  
3. 针对查询类别的“添加属性限制”从属性上约束查询类;  
4. 对于非查询类别“类别实例化”可以输入实例来限定该类别;  
5. “仅保留该类别”将删除选中类别外所有类别和联系;

显示查询信息

```
graph TD; A(副教授 as QueryVariable) -- 授课 --> B(课程);
```

# User-defined graphics mode

新的查询 new    清空查询 clear    递交查询 query

2. 针对所有类别的“关联到其它类别”可以建立类别间的关系；  
3. 针对查询类别的“添加属性限制”从属性上约束查询类；  
4. 对于非查询类别“类别实例化”可以输入实例来限定该类别；  
5. “仅保留该类别”将删除选中类别外所有类别和联系；  
6. “删除选中类别”将删除选中类别及关联的联系；

显示查询信息 display

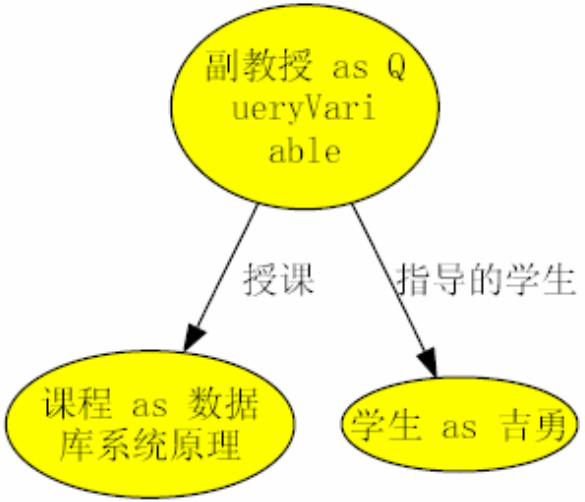
```
graph TD; A(副教授 as QueryVariable) -- 授课 --> B(课程 as 数据库系统原理)
```

# User-defined graphics mode

新的查询 new    清空查询 clear    递交查询 query

2. 针对所有类别的“关联到其它类别”可以建立类别间的关系；  
3. 针对查询类别的“添加属性限制”从属性上约束查询类；  
4. 对于非查询类别“类别实例化”可以输入实例来限定该类；  
5. “仅保留该类别”将删除选中类别外所有类别和联系；  
6. “删除选中类别”将删除选中类别及关联的联系；

显示查询信息 display



```
graph TD; A(副教授 as QueryVariable) -- 授课 --> B(课程 as 数据库系统原理); A -- 指导的学生 --> C(学生 as 吉勇);
```

# User-defined graphics mode

新的查询  清空查询  递交查询

2. 针对所有类别的“关联到其它类别”可以建立类别间的关系；  
3. 针对查询类别的“添加属性限制”从属性上约束查询类；  
4. 对于非查询类别“类别实例化”可以输入实例来限定该类别；  
5. “仅保留该类别”将删除选中类别外所有类别和联系；  
6. “删除选中类别”将删除选中类别及关联的联系；

显示查询信息

```
graph TD; A("副教授 as QueryVariable") -- 授课 --> B("课程 as 数据库系统原理"); A -- 指导的学生 --> C("学生 as 吉勇"); A -- 籍贯 --> D("as 湖北");
```

# User-defined graphics mode

SmartCH

定制搜索

搜索

[返回首页](#) [图形化](#)

共有22项符合 **副教授** 的查询结果 (搜索用时: 0.656秒)

## 《李瑞轩》

[华中科技大学智能与分布计算实验室](#)

李瑞轩 李瑞轩 李瑞轩, 文坤梅 卢正鼎, 李瑞轩...

<http://idc.hust.edu.cn:7001/idc/resDirection/resDirectionList.do?pageCount=1>

[华中科技大学智能与分布计算实验室](#)

李瑞轩 李瑞轩 李瑞轩, 文坤梅 卢正鼎, 李瑞轩...

<http://idc.hust.edu.cn:7001/idc/resDirection/resDirectionList.do>

[华中科技大学智能与分布计算实验室](#)

李瑞轩, 卢正鼎. 李瑞轩, 杨曼红, 王俊....

<http://idc.hust.edu.cn:7001/idc/proBook/proBookList.do>

[华中科技大学智能与分布计算实验室](#)

彭俊杰, 卢正鼎, 李瑞轩... 胡和平, 黄保华, 姚寒冰, 卢正鼎, 李瑞轩... 唐卓, 卢正鼎, 李瑞轩, 霍晓丽, 陈媛玲, 李瑞轩, 李瑞轩, 赵战西, 王治纲, 卢正鼎, 唐卓, 卢正鼎, 李瑞轩, 李红平... 凌贺飞, 卢正鼎, 邹...

李瑞轩详细资料:

指导的学生: 柯峰

发表著作: 一种P2P环境中基于语义的Web服务组合方法

指导的学生: TerefeMuleta

授课: 数据库系统原理

研究方向: 分布计算; 分布式异构系统, 对等计算, 边缘计算

发表著作: 在Java2环境中实现可插入的认证及访问控制

type: 副教授

电子邮件: rdi@public.wh.hb.cn

出生年月: 1974

发表著作: paper004

发表著作: paper006

授课: 系统集成与中间件

研究方向: 数据管理; 异构数据集成, Web数据管理, 语义网与本体论

授课: 数据库系统分析与设计

研究方向: 系统安全; 分布式系统安全

发表著作: paper003

发表著作: paper004

# Semantic Association search

MyIE2 - [Semantic Association Search(SAS)]

文件(F) 编辑(E) 查看(V) 收藏(A) 快捷组(G) 选项(O) 工具(T) 窗口(W) 帮助(H)

地址 http://211.69.192.118:8085/sas/search.jsp

Semantic Association Search(SAS)

Length weight:   Long Association Prior  
Context weight:  Node In&Out weight:

Start Node:  孙小林

End Node:  卢正鼎

Show Explanation

Please select the classes you want to emphasize:

- http://www.owl-ontologies.com/unnamed.owl#教授
- http://www.owl-ontologies.com/unnamed.owl#副教授
- http://www.owl-ontologies.com/unnamed.owl#著作
- http://www.owl-ontologies.com/unnamed.owl#讲师
- http://www.owl-ontologies.com/unnamed.owl#书籍
- http://www.owl-ontologies.com/unnamed.owl#项目
- http://www.owl-ontologies.com/unnamed.owl#实验室
- http://www.owl-ontologies.com/unnamed.owl#硕士



# Semantic Association search

Time: 4405 ms. Existing 5009 semantic association relationship.

<http://www.owl-ontologies.com/unnamed.owl#孙小林> and <http://www.owl-ontologies.com/unnamed.owl#卢正鼎> have following association relationship!

1. 孙小林 指导教师 卢正鼎  
Total Value: 0.42523062
2. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 指导教师 卢正鼎  
Total Value: 0.24529517
3. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 发表著作 paper005 作者 卢正鼎  
Total Value: 0.20862354
4. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 发表著作 paper005 作者 王治纲 指导教师 卢正鼎  
Total Value: 0.184663
5. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 发表著作 paper005 作者 李瑞轩 发表著作 在Java2环境中实现可插入的认证及访问控制 作者 卢正鼎  
Total Value: 0.17508751
6. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 发表著作 paper005 作者 李瑞轩 发表著作 paper003 作者 卢正鼎  
Total Value: 0.17483343
7. 孙小林 发表著作 基于本体的多域访问控制策略集成研究 作者 文坤梅 发表著作 paper005 作者 李瑞轩 发表著作 一种Peer-to-Peer环境下的分布式访问控制模型 作者 卢正鼎  
Total Value: 0.17481378

---

# Conclusion

- User-defined graphics mode can inerrably ensure the semantic information which hides in user's query and improve the precision.
- Smarth also can find out the association relationship between two entities.

---

# Future work

- Large scale ontology data
- Better efficiency and speed
- More handy user-defined graphic mode

---

**Thank you!**